■主な製品・機能

全文検索スウィート	Bibliotheca21 Version 2 Standard	Windows®ファイルシステム、Webサーバ、Notes／Domino、RDBなどの情報を一括検索する全文検索パッケージ
高速全文検索エンジン	Bibliotheca2 TextSearch Server Version 2	検索／DB管理／ユティリティ
	Bibliotheca2 TextSearch Server Library Version 2	Webサーバ向け実行環境
	Bibliotheca2 TextSearch Gateway Version 2	Webインタフェース
ミッションクリティカル業務向け全文検索エンジン＊1	HiRDB Text Search Plug-in Version 8	HiRDB Version 7用日本語全文検索プラグイン
	HiRDB Text Search Plug-in Conceptual Extension Version 7	自然文検索機能
	Text Search Enhancer for English	英文検索機能
	Text Search Filter Library Version 3	ヒット位置情報取得ライブラリ
文書管理用検索オプション＊2	DocumentBroker Collector	文書の収集、差分収集、分割収集、文書登録
	DocumentBroker Collector文書収集オプション	Webサーバ、Groupmax Document Manager、Lotus Domino（Notes）、RDB（HiRDB、Microsoft® SQL Server™、ORACLE）からの各収集オプション
	DocumentBroker Integrated Search Suite for J Version 2	検索ポータル構築用パッケージ
各製品共通オプション	Text Search Dictionary	同義語辞書

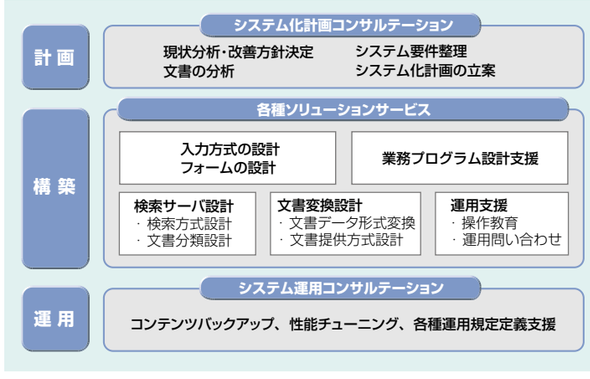
＜稼働環境＞Windows Server® 2008、Windows Server® 2003、Windows Vista®、Windows® XP Professional、Windows® 2000、HP-UX、AIX 5L、Solaris、Linux
各製品のOSサポート状況については、「日立高速全文検索シリーズ」のホームページをご参照くださるか、別途、お問い合わせください。

＊1 前提製品として、HiRDBサーバおよびその前提・関連プログラムが必要になります。　＊2 前提製品として、DocumentBrokerサーバおよびその前提・関連プログラムが必要になります。

ソリューションサービス

お客さまのニーズに合わせた最適なシステムの構築をサポート。

文書管理/文書検索システムの適用業務について、日立の専門技術者が、計画から構築、運用までを一貫してサポート。豊富な構築実績で培ったノウハウと先進の技術を結集し、各種コンサルテーションやソリューションサービスを組み合わせて、お客さまの業務形態に最適なシステムを構築します。



・AIXは、米国における米国International Business Machines Corp.の登録商標です。
・HP-UXは、米国Hewlett-Packard Companyのオペレーティングシステムの名称です。
・Linuxは、Linus Torvaldsの米国およびその他の国における登録商標あるいは商標です。
・Lotus、Lotus Domino、Lotus Notesは、IBM Corporationの商標です。
・Microsoft、Windows、Windows NT、Windows Server、Windows Vista、Visual Basic、Visual C++は、米国Microsoft Corporationの米国およびその他の国における登録商標または商標です。
・Microsoft Word、Microsoft Excel、Microsoft SQL Serverは、米国Microsoft Corporationの商品名称です。
・ORACLEは、米国Oracle Corporationの登録商標です。
・Solarisは、米国およびその他の国におけるSun Microsystems, Inc.の商標または登録商標です。
・SQL Serverは、米国法人Sybase, Inc.の商標です。
・一太郎は、株式会社ジャストシステムの登録商標です。
・その他記載の会社名、製品名は、それぞれの会社の商標もしくは登録商標です。

●カタログに記載の仕様は、製品の改良などのため予告なく変更することがあります。　●製品の色は印刷されたものですので、実際の製品の色調と異なる場合があります。
●本製品を輸出される場合には、外国為替及び外国貿易法並びに米国の輸出管理関連法規などの規制をご確認のうえ、必要な手続きをお取りください。
なお、ご不明な場合は、弊社担当営業にお問い合わせください。

製品に関する詳細・お問い合わせは下記へ

■ 製品情報サイト
http://www.hitachi.co.jp/soft/textsearch/
■ インターネットでのお問い合わせは
http://www.hitachi.co.jp/soft/ask/
■ 電話でのお問い合わせは**HMCC**（日立オープンミドルウェア 問い合わせセンター）へ
0120-55-0504 利用時間 9:00～12:00、13:00～17:00（土・日・祝日・弊社休日を除く）

株式会社 日立製作所　情報・通信グループ　ソフトウェア事業部

CA-371Q　2009.9
Printed in Japan(H)

日立高速全文検索シリーズ

HITACHI
Inspire the Next

知を活かすビジネスはここから生まれる。

日立高速全文検索シリーズ

Bibliotheca21 ビブリオテカ21
Bibliotheca2 ビブリオテカ2 TextSearch
HiRDB Text Search Plug-in
DocumentBroker Collector/Integrated Search Suite for J

uVALUE with Harmonious Computing

本カタログはAdobe社Acrobatにより制作したPDFカタログです。
All Rights Reserved,Copyright © 2009,Hitachi,Ltd.

選り抜かれた検索システムがここに集結。
今こそ、価値ある情報をビジネスに活かす時。

企業内外に蓄積される膨大な情報。そこからつねに必要な情報を瞬時に手に入れることができれば、業務の効率化はもちろん、新しいビジネスを生み出すことも夢ではありません。そこで日立では、高速で漏れのない全文検索を実現する多彩な製品群を「日立高速全文検索シリーズ」として提供。個人利用から部門レベルの小規模システム、社内ポータルやミッションクリティカル業務などの大規模システムまで適応し、お客さまの規模や用途に合わせて適切な検索システムをご提供いたします。

日立高速全文検索シリーズ

Bibliotheca21
Bibliotheca2 TextSearch
HiRDB Text Search Plug-in
DocumentBroker Collector/Integrated Search Suite for J

部門内のネット環境に埋もれがちな情報を誰でも簡単に検索できるようにしたい

全文検索スウィート
Bibliotheca21 ビブリオテカ21

簡単な導入で収集から検索までのシステムを構築し、企業内の情報共有を促進します。

業務システム内に蓄積する情報から必要なものを的確に取り出して活用したい

高速全文検索エンジン
Bibliotheca2 ビブリオテカ TextSearch

業務システムに高速な全文検索エンジンを組み込み、さまざまな検索条件で多角的に検索できます。

eビジネスで扱う膨大なテキスト情報の検索システムがほしい

ミッションクリティカル業務向け全文検索エンジン
HiRDB Text Search Plug-in

1,000万件単位の文書の全文を対象とした検索サービスを実現します。

複数のシステムやデータベースに分散している情報をシームレスに収集して活用したい

文書管理用検索オプション
DocumentBroker Collector/ Integrated Search Suite for J

情報の収集とデータベースに蓄積する情報の横断検索を実現します。

高度で柔軟な全文検索技術が、効率的で使いやすい情報環境を支えていきます。

価値ある情報を迅速にビジネスへ活かす。そのためには、文書のフォーマットや管理環境の違いにとらわれずに文書を「集め」、さらに収集した文書の中から目的の文書を効率的かつ的確に「探し出す」ための機能が必要です。こうした要件にお応えするため、日立は豊富で高度な検索・収集機能を提供し、お客さまのビジネス環境における情報資産活用のサイクルを具現化します。

検　索

●大量文書を漏れなく高速検索

検索アルゴリズムとして日立が誇るインクリメンタルn-gramインデックス方式を採用。1サーバで100万件を超える文書を漏れなく高速に検索することが可能です。

●多彩な検索機能をサポート

文書で検索できる自然文検索(概念検索)をはじめ、キーワード検索（全文検索）や表記のゆれを吸収できる同義語・異表記展開検索など、さまざまな検索手段を提供しています。

●辞書メンテナンスフリー

検索用インデックスの作成に辞書を必要としないため、辞書メンテナンス無しで、新語・造語に対応したキーワード検索、自然文検索を行えます。

収　集

●さまざまな情報環境から収集

Webサーバ、Windows®ファイルシステム、グループウェア、リレーショナルデータベースなどのネットワーク上に分散している文書を、収集先の環境に合わせてシームレスに収集します。

●多様な文書に対応

Microsoft® Wordや、Microsoft® Excel、一太郎、PDF、HTML、テキストなどの多様な形式の文書からテキスト情報を抽出し、XML形式に変換して取り扱うため、フォーマットの違いにとらわれることなく収集できます。

●効率的に一括収集

文書の収集範囲、収集する文書数、処理時間、収集目的などによって収集条件を選択し、効率的にきめ細やかな収集が行えます。

※収集機能に関しては全文検索スウィート「Bibliotheca21」、文書管理用検索オプション「DocumentBroker Collector」のみのサポートになっております。

日立高速全文検索シリーズ

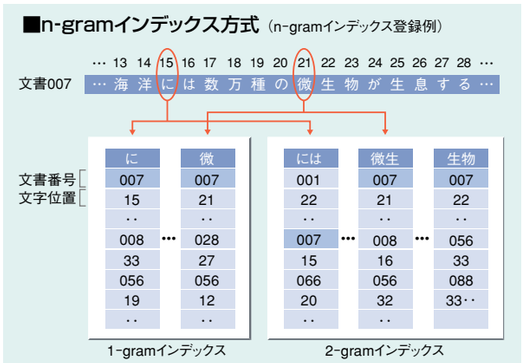
Bibliotheca21、Bibliotheca2 TextSearch、HiRDB Text Search Plug-in、DocumentBroker Collector/Integrated Search Suite for J

永年の実績から生まれた先進の検索・収集技術を結集させています。

検　索

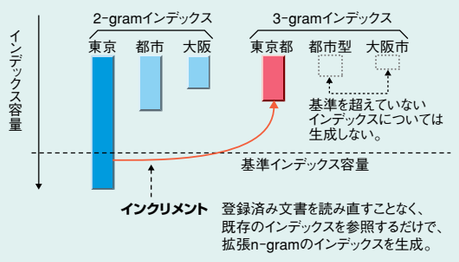
日立が誇るインクリメンタル n-gramインデックス方式

文章を単語で区切らず、n文字の文字列そのものをインデックスにして、検索を実現するn-gramインデックス方式を採用。単語の情報を格納した辞書を必要とすることなく、漏れのない高速全文検索を実現し、新語・造語にも即時に対応できます。



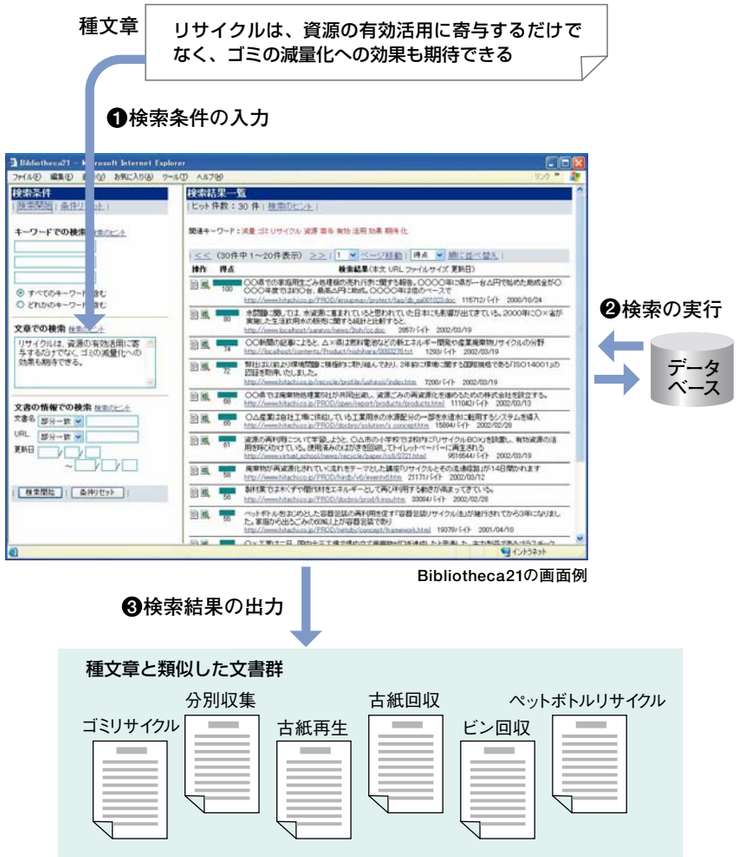
また、一般的にn-gramインデックス方式は、n-gram長を拡張すると（たとえば2-gramのインデックスに加えて、3-gramインデックスを生成）、検索性能が向上する一方でインデックスの生成には時間がかかります。しかし、日立が誇るインクリメンタルn-gramインデックス方式なら、インデックス容量およびインデックスの生成にかかる時間の増加を抑えながらインデックスのn-gram長を拡張し、検索スピードを向上させることができます。

検索に時間のかかるn-gramに対し、n-gram長を拡張したインデックスを生成 → 総インデックス容量を抑えつつ検索性能を確保



文章を入力して検索できる自然文検索をサポート

たとえば報告書を作成する際に関連記事や類似報告書を探したいといった場合、任意の文章や文字列（種文章）を手がかりにして探し、その文章や文字列に近い内容をもつ文書を検索できます。キーワードや検索式を指定することが困難な場合は、文章を検索条件にすると便利です。



Bibliotheca21の画面例

基本検索機能ラインアップ

キーワード検索
文書中のキーワードを検索条件に指定して検索します。また、2つ以上のキーワードを指定してそれらがすべて含まれる文書（AND条件）、またはそれらのうちどれかが含まれる文書（OR条件）を検索することもできます。

自然文検索
具体的なキーワードが思いつかない場合、検索したい内容を入力した文章や文字列（種文章）を検索条件に指定して、内容（概念）が類似した文書を検索します。検索条件として指定した文章から抽出したその文章を特徴付ける単語（関連キーワード）を使用して検索します。

複合検索
文書中のキーワード検索条件、自然文検索条件（種文章）、プロパティ検索条件を、同時に指定して検索します。キーワードと種文章などを同時に指定して検索することで、より目的に近い検索結果を得ることができます。

関連文書検索
検索結果の中から検索条件（種文章）とする文書を選択して、その文書と内容が類似した文書（関連文書）を検索します。最初の検索条件で目的の文書が検索できなかった場合に、検索結果から目的の文書の内容に似た文書を選択して検索できるため、目的の文書がヒットしやすくなります。

同義語展開検索
同義辞書語を用いて、指定したキーワードと同じ意味を持つ語（例：アメリカ、米国、USA）も同時に検索できます。なお、同義語辞書はお客様に作成いただく必要があります。また、オプションとして約11,000語を収録した同義語辞書の提供も行っております。

異表記展開検索
同じ意味を持ちながら、アルファベットの小文字・大文字（例：ski、SKI、Ski）、カタカナ（例：バイオリン、ヴァイオリン）、および全角・半角文字などの表記のゆれによる異表記も同時に検索できます。
※カタカナ異表記のルールは、国語審議会の答申に基づき、日立でルール化したものを採用しています。

スコア検索
検索する際に指定した検索条件を満たす度合いによって、検索結果の各文書にスコア（得点）を付け、得点順にソートして検索結果を並べて表示します。

収　集

効率的できめ細やかな収集が可能

文書の収集先に合わせて、URLやファイルシステム名、グループウェアまたはリレーショナルデータベースのサーバ名などを指定し、容易に文書の収集範囲を定義できます。また、収集する文書数、収集にかかる時間、収集目的によって収集条件を選択し、効率的できめ細やかな収集ができます。

新規収集
収集対象となっているすべての文書を収集。すでに収集したことのある文書を再度収集することもできます。

差分収集
収集対象になっている文書のうち、前回収集時から追加および更新のあった文書だけを収集します。

分割収集
収集対象となっている文書を数回に分けて収集。収集対象となる文書数が多い場合に分割収集を使うことで、1回に収集する文書数および収集にかかる時間を調整できます。

差分分割収集
追加および更新のあった文書だけを数回に分けて収集。差分を分割して収集している間に追加された文書は、収集する時点の状態で収集できます。

条件式収集
文書の属性に対する検索条件を記述して、条件に当てはまる文書を収集。たとえば「2003年2月14日以降に更新があった文書だけを収集する」などの収集が可能です。また、文書を収集しないで条件にあてはまる文書の件数だけを取得することもできます。

※収集機能に関しては全文検索スウィート「Bibliotheca21」、文書管理用検索オプション「DocumentBroker Collector」のみのサポートになっております。

高度な検索システムを容易に構築し、分散する文書を一括検索。

全文検索スウィート Bibliotheca21 ビブリオテカ21

インターネット/イントラネット、共有ドライブ、RDB、LotusNotes/Dominoなど、さまざまなネットワーク環境に分散する文書を、Webブラウザを使って一括検索する検索パッケージ。高度なキーワード検索（全文検索）や自然文検索（概念検索）を駆使し、目的の文書をすばやく的確に抽出することが可能です。また、導入をはじめ、文書収集の運用やデータベースの作成・削除、最適化も容易です。

規模に合わせて最適なシステムをお選びいただけます。

Bibliotheca21 Standard（中規模システム向け）

Windows®ファイルシステム、インターネットホームページ、リレーショナルデータベースおよびLotusNotes/Domino対応の全文検索パッケージ。登録文書件数に制限はなく（推奨登録文書件数は300万件）、オプション製品により機能拡張が可能です。文書管理を必要としない中規模の検索システムとして適切です。

利用者のメリット

●文書の格納場所や文書の形式を問わずに一括検索

さまざまなネットワーク環境やローカルディスクなどに蓄積する文書を収集し、検索用のデータベースに登録。WebおよびWindows®のネットワーク環境であれば、格納場所や文書形式を問わずに一括して検索できます。

●豊富な検索手段で目的の文書を的確に検索

キーワード検索、自然文検索、属性検索が行えます。また、これらの機能を組み合わせて検索したり、検索結果の中から選んだ文書を元に検索して、関連した情報をもつ文書を探すこともできます。

管理者のメリット

●スムーズに文書検索システムを構築

収集環境定義ウィザードとテキストデータベースの作成を支援するデータベースユーティリティの2種類のGUIを提供し、文書検索システムの構築における複雑な手順を簡易化し、スムーズに導入できます。

●テンプレートを利用してスピーディーに運用開始

Webブラウザから文書の検索が行えます。文書検索用の画面を検索テンプレートとして用意し、Webサーバのネットワーク環境を設定するだけで、文書検索システムの運用をスピーディーに開始できます。

●利用環境に合わせた効率的なシステム運用

最新の文書を対象に収集を行う差分収集や、文書の収集を一定の時間や文書数ごとに数回に分けて実行する分割収集といった機能を利用することで、効率的かつ柔軟にシステム運用が行えます。

●オプション製品による機能拡張

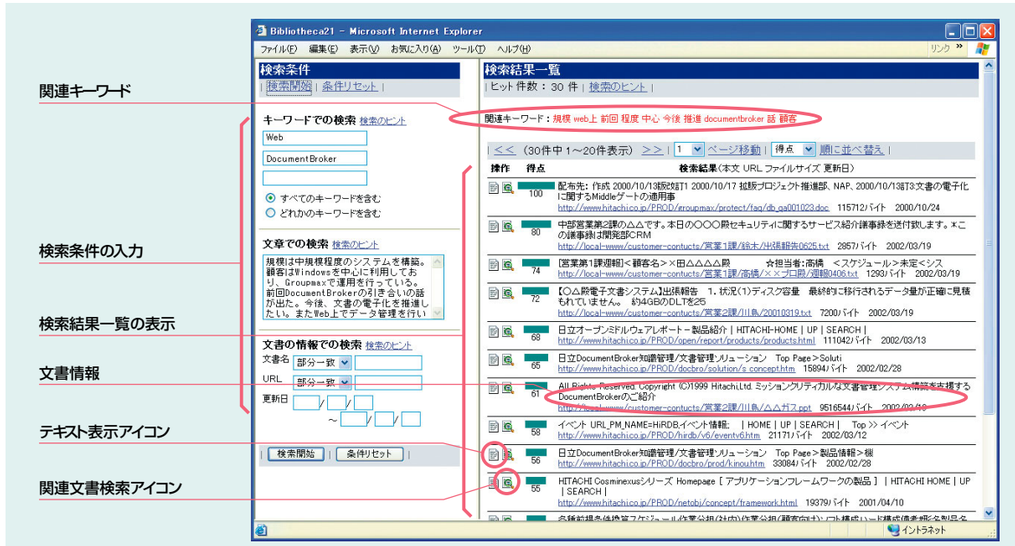
予約検索

定期的に特定のキーワードを含む文書を検索した結果を、メールで受信可能。最新の情報を容易に把握できるようになります。

画面カスタマイズ

検索画面カスタマイズや検索クライアントアプリケーションを作成するためのCOMのインタフェース群を提供します。Visual C++®、Visual Basic®、ASPなど、さまざまな開発環境で利用することができます。

COM：Component Object Model
ASP：Active Server Pages



適用例

社内イントラネット上のファイルサーバで共有している文書を一括検索

導入目的

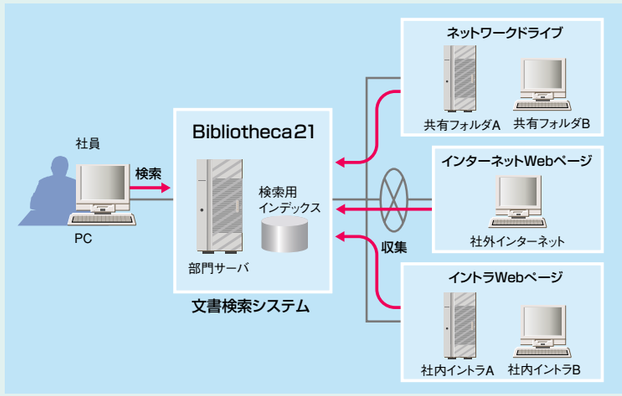
文書を部門内のイントラネット上のファイルサーバに保管して、社員が簡単に検索できるようにしたい。

システムの特長

- ●社内通達やインフラ情報、ノウハウ集など、さまざまな場所にある文書を一括検索。
- ●文書検索用のデータベースの運用が不要。
- ●収集環境定義のウィザードの画面の支持に従うだけで文書収集環境を簡単に設定可能。

効果

- ●さまざまな場所から必要な文書をWebブラウザ上ですばやく探し出すことができるようになった。
- ●柔軟な文書収集機能を利用してスムーズに運用できるため、管理者の手間が不要になった。



世界最高レベルの検索性能を誇る検索エンジンをあなたのシステムで。

高速全文検索エンジン Bibliotheca2 ビブリオテカ2 TextSearch

100万文書の情報を1秒以内*で検索できる世界最高レベルの検索エンジン。既存の業務システムへの全文検索機能の組み込みを実現し、多彩な検索方法で求める情報を高速に検索できます。

＊新聞記事換算、弊社環境測定。

●用途に合わせてきめの細かい柔軟な検索が可能

新しい言葉でも言葉の一部からでも、高速かつ正確に必要な文書を探すことができます。また、検索結果をスコア順に表示するスコア検索、さらに検索結果のなかでヒットした位置をハイライト表示することもできます。

● 構造指定検索のサポート

XMLなどで書かれた構造化文書（見出しや段落などの構造を明確に定義した文書）を登録すると、文書中の構造を指定して検索可能。より正確に必要な文書を探し出せます。

XML：eXtensible Markup Language

適用例 図書情報パッケージに高速全文検索エンジンを搭載

導入目的

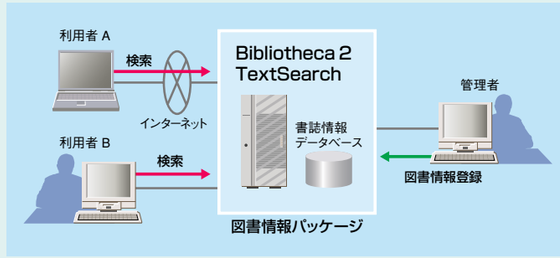
日々増加する膨大な量の図書情報の中から、誰もがいつでも必要な情報を簡単に得られるようにしたい。

システムの特長

- ●資料種別や新刊・所蔵・発注情報の指定／混在検索、論理検索（AND・OR・NOT）、完全一致検索、前方一致検索など、さまざまな検索方法に対応。

効果

- ●利用者が自宅や会社に居ながらにして、いつでもWebブラウザからアクセスして目的の資料をすばやく検索できるようになった。



企業内の大量文書をデータベースに蓄積し、高速検索を 実現。

ミッションクリティカル業務向け
全文検索エンジン

HiRDB Text Search Plug-in

データベースへの文書登録や、特定の文字列を含む文書の高速検索に向けた日本語全文検索プラグインとして、スケーラブルデータベース「HiRDB」に組み込んで使用する、大規模・ミッションクリティカル業務に最適な全文検索エンジンです。全文検索用テキストデータとコード系データを同一トランザクションで操作したり、大量データをパラレルに検索することができます。

●日本語の検索にぴったりの条件指定

単なるテキストの全文検索にとどまらない豊富な検索方法を利用できます。XML*1などの文書の構造を意識した検索といった高度な機能も提供しています。

*1 XML文書を利用する際は、XML解析用の別製品が必要となります。

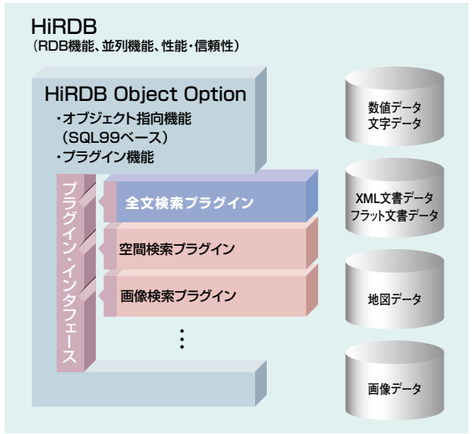
●Unicode対応*2

テキストデータを登録する際に指定できる文字コードとして、Shift JISコード、EUCコードに加え、Unicode（UTF-8*3）をサポートしています。

*2 対象はXML文書のみ　*3 MS-UnicodeおよびJISX0221に対応

●SQLを利用して文書や文字列を容易に検索

HiRDB（+Object Option）は、SQL99*4をベースにして、文書、地図、画像といったデジタルコンテンツやその操作



機能を容易に取り込めるプラグインアーキテクチャを採用しています。文書や文字列に対する複雑なデータ操作もSQLで定義し、SQLを利用して検索することができます。

*4 SQL1992の後継国際標準規格（1999年制定）

●用途に応じて選択可能な拡張機能

自然文検索

キーワードの代わりに文書や文字列を検索条件に指定して、それに類似した文書を探すことのできる自然文検索が可能です。

英文検索

英文検索用のインデックスを作成でき、英語特有の半角アルファベット異表記、英単語派生表記、英単語語尾表記を考慮して英単語を検索できます。

ヒット位置情報取得用ライブラリ

検索条件として指定したキーワードの出現している位置を取得するヒット位置情報取得用のアプリケーションを作成するためのライブラリ機能を提供します。

適用例　電子カタログ検索によるWeb受発注システム

導入目的

膨大なカタログデータの中から必要な製品情報を高速かつ柔軟に検索できるWeb受発注システムを構築したい。

システムの特長

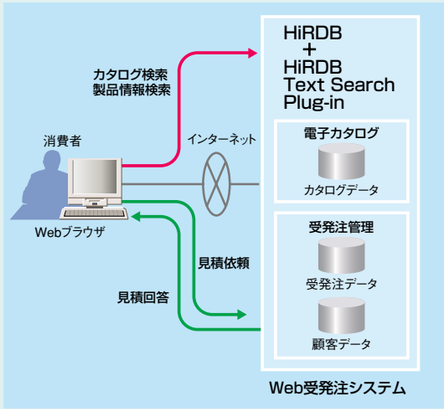
●テキストデータとコード系データを同一トランザクションで高速検索。カタログ中のテキスト情報を対象とした「キーワード検索」、各カタログの属性情報を検索キーとした「属性検索」を実現。またそれらを組み合わせた検索により、膨大な商品カタログのアイテム情報の中から効率的かつ的確に検索可能。

●拡張性に富んだHiRDBとの連携により、カタログの全アイテムをチェックするような全文検索でも高速なレスポンスを実現。

効果

●柔軟で使い勝手の良いWeb受発注システムが構築できた。

●データ量やユーザー数が増大しても、複数サーバによる並列化により容易にシステムを拡張できる。



さまざまな環境に分散する多様な形式の文書を一括収集 し、高度な検索が可能。

文書管理用
検索オプション

DocumentBroker Collector/ Integrated Search Suite for J

文書管理基盤製品である「DocumentBroker」と連携する文書管理用検索オプションです。さまざまな環境に分散している多様なフォーマットの文書からテキストデータおよび属性情報を抽出し、XML形式に変換して一括収集します。また多彩な検索機能を組み合わせて文書を多角的な視点から検索し、効率的かつ的確に探し出すことができます。

●さまざまな環境に分散した文書の一括収集

多様な文書、大量の文書を統一的に収集

Webサーバ、Windows®ファイルシステム、グループウェア、リレーショナルデータベースなどの環境に格納されている文書（テキスト情報）からテキストデータを収集可能。同時にMicrosoft Word形式、PDF形式やHTML形式など、著名な文書形式から文書の属性（作成者、作成日）などを自動抽出してDocumentBrokerのデータベースへ蓄積します。

●効率的かつ的確に文書を検索

高度な検索機能

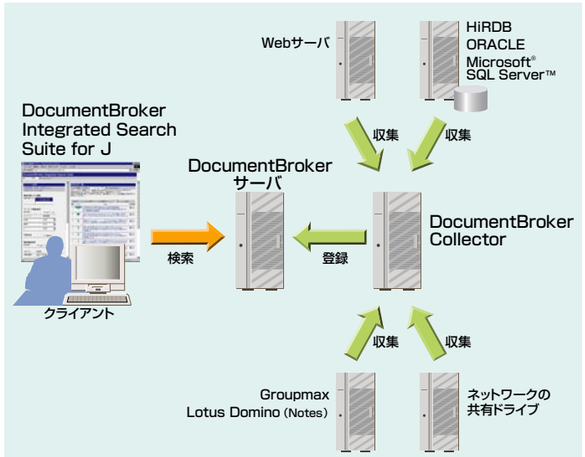
キーワード検索、自然文検索、属性検索など多彩な検索機能を提供しています。

利用形態に適した検索画面を容易に作成

さまざまな検索機能を複合して利用できる検索テンプレートを標準で提供しています。このため検索画面を最初から作成する必要はありません。また、この検索テンプレートを利用形態に合わせてカスタマイズすることも可能です。

検索専用のポートレット機能を提供

Webアプリケーションサーバ「Cosminexus」を使用して構築したシステム上に、容易に検索機能を組み込める検索専用のポートレット機能を提供しています。



システム構築および運用が容易

文書の収集先や収集する文書属性など、システム構築を定義した各種定義ファイルのサンプルを提供しています。このため、より少ないカスタマイズでシステムの構築が可能。また、収集する範囲や文書が多量な場合、追加および更新された文書だけを収集する差分収集機能、収集を数回に分けて行える分割収集機能を提供しています。

適用例　Lotus Notes/Domino上の営業支援情報の横断検索

導入目的

複数のLotus Notes/Dominoに分散しているカタログ、取扱説明書などの製品情報やFAQ、クレーム情報といった営業支援用の共有情報から、必要な文書を的確に探し出したい。

システムの特長

●複数のノーツデータベースにある情報の収集、横断検索を実現。

●キーワード検索、自然文検索に加え、ノーツデータベース間の共通属性情報（製品コード、製品名称、製品分類、タイトル、登録日など）で検索する機能をサポート。

●Lotus Domino（Notes）の他に、RDBやWeb／ファイルサーバの情報も収集が可能な拡張性の高いシステムを提供。

効果

●複数のノーツデータベースに分散していた営業支援情報を横断して検索できるようになった。

●RDBに蓄積されている情報も検索対象として追加することが可能となった。